Speech
Etymology
Lemma
Idioms
Glossary
NLP
Corpora
Dictionary
Meaning
Word
Lexicon
Definition
Pronunciation
Headword
Entry
Lexicology
Lexical Resources
Dictionary Use

λ EURALEX XIX

Congress of the European Association for Lexicography

Lexicography for inclusion

7-9 September 2021
Virtual

www.euralex2020.gr

Proceedings Book
Volume 2

Edited by Zoe Gavriilidou, Lydia Mitits, Spyros Kiosses

2021 Edition

Using verb patterns to find recurrent metaphors in corpus

Renau I.

Pontificia Universidad Católica de Valparaíso
irene.renau@gmail.com

Abstract

In this study, we examine the possibility of finding regularities in combinations of verb patterns, and if these regularities can be used to find recurrent metaphors in discourse. As the source of the data, we used Verbario, a database of 227 Spanish verbs that were annotated with the Corpus Pattern Analysis technique (Hanks 2004, 2013). We restricted our analysis to transitive patterns in order to have identical syntactic structures and be able to focus our analysis to semantic types only. Given a verb pattern such as *[[Humano]] guarda [[Objeto Físico]] ([[Human]] keeps [[Physical Object]]),* the base pattern is *[[Human]] ~ [[Physical Object]],* a syntacto-semantic structure that can be found also in verbs other than *guardar.* 177 verbs from the database (78%) had 2 or more transitive structures and were included in the study. Results show how a small number of semantic types and combinations of verb patterns are linked to most of the verbs. Additionally, many pairs of base patterns are connected to each other through metaphors. The study is of interest for lexicographic tasks involving corpus analysis and is a contribution to corpus-based studies of metaphor.

Keywords: Corpus Pattern Analysis; metaphor; polysemy; semantic type; Spanish

1   Introduction

Metaphors have often been classified as irregular polysemy, especially in contrast with metonymy, which is usually linked to regular polysemy (Apresjan 1974). While polysemy by metonymy is more systematic and predictable, polysemy by metaphor is more idiosyncratic and accidental, and it is not predictable. However, the theory of metaphor (Lakoff & Johnson 1980) postulates that metaphor, as well as metonymy, are cognitive (not only linguistic) resources that people exploit to categorize the world and communicate: thus, metaphors have to hold a certain degree of regularity in order to be used and understood. A metaphor such as "Humans are Machines" is not predictable nor systematic, but it lays beyond many linguistic expressions such as *My mind does not work well today, I have too many memories in my hard drive,* etc. This is precisely what conceptual metaphors are. The same happens with conceptual metonymies: a metonymy such as "Plant Part for the Plant" is a cognitive resource that can be exploited, but we cannot predict when a plant part is going to be labeled with the name of the plant, and the same happens with colors/flours, products/plants, etc. (Renau 2021). Hence, there is probably not a sharp distinction between regular polysemy, irregular polysemy and homonymy, but a gradual distinction (Moldovan 2019).

In this study, we propose a method to find regularities in combinations of verb patterns which could be used to find recurrent metaphors in discourse. We take the concept of verb pattern from Hanks' Corpus Pattern Analysis, CPA (Hanks 2004, 2013, among others). The author (Hanks 2004: 87) states that word meaning is associated to "syntagmatic patterns with which words in use are associated". Thus, in real-life discourse, these patterns, consisting of the basic valency structure and other semantic and syntactic features, are the ones carrying the meaning of the verb, and not the verb in isolation. For our investigation, we use a database of Spanish verbs (Renau et al. 2019) that were annotated following the CPA principles. Given two verb patterns, we examine the possibility of finding an equivalent pair of patterns in other verbs. Observe the following examples:[1]

(1)   Verb *ensanchar* 'to widen'
        Pattern 1: [[Humano]] ensanchar [[Objeto Físico]] *([[Human]] widens [[Physical Object]])*
        Example: "En verano ensancharon el camino" *('In the summer they widened the road').*

        Pattern 2: [[Humano]] ensanchar [[Entidad Abstracta]] *([[Human]] widens [[Abstract Entity]])*
        Example: "[Ustedes] sabrán preservar y ensanchar nuestra armónica convivencia" *('You will know how to preserve and widen our harmonious coexistence').*

(2)   Verb *guardar* 'to keep'
        Pattern 1: [[Humano]] guarda [[Objeto Físico]] *([[Human]] keeps [[Physical Object]])*
        Example: "Compró el libro y lo guardó en el bolsillo del abrigo" *('[She/he] bought the book and kept it in the pocket of her/his coat').*

---

[1] See the complete, original analysis of all Spanish patterns in the Verbario database (http://www.verbario.com). For clarity's sake, in the paper we simplified some of the patterns. All examples are from the EsTenTen corpus, available in the Sketch Engine (Kilgarriff et al. 2014).

Pattern 2: [[Humano]] guarda [[Entidad Abstracta]] *([[Human]] keeps [[Abstract Entity]])*
Example: "Ya están grandes y saben lo que implica guardar un secreto" *('They are grown up enough to know what it means to keep a secret').*

In (1), pattern 1 and pattern 2 of *ensanchar* 'to widen' are linked by a metaphor whereby an [[Abstract Entity]] is categorized as a [[Physical Object]] that can be "widen". We could formulate this metaphor as "Abstract Entities are Physical Objects", which is one of the most used conceptual metaphors. The interesting is that we can find the same relation in other verbs, such as *guardar* 'keep' in (2). Pattern 1 and pattern 2 of both verbs share the same semantic types in the argument structure. These common patterns could be expressed as follows:

(3)   *Base pattern 1:* HUMAN ~ PHYSICAL OBJECT
        Base pattern 2: HUMAN ~ ABSTRACT ENTITY

We call *base pattern* to the abstract pattern consisting of the semantic types and argument structure only, without the verbs and with no link to any specific meaning. (Base patterns are indicated with small capitals.)
Following this rationale, we wonder if we can find similar associations such as the ones shown in examples (1) and (2) by extracting base patterns such as (3) from the CPA patterns that we already have in our database. Do all these associations have a metaphorical nature? Are they relatively stable? Having a base pattern X, can we predict that a base pattern Y is going to appear in the same verb?
This study is of interest for lexicography by contributing to automatic techniques to interrogate a corpus for lexicographic purposes (Kosem 2016). Specifically, it could be of help regarding the so-called "pre-lexicographic" or "preliminary" tasks of the dictionary-making process (Atkins & Rundell 2008, Hartmann 2001), particularly for collecting and analyzing corpus data. Corpus analysis is necessary as the empirical basis of the information offered in a dictionary, but it is still a very time-consuming and complicated task which has to be executed manually to a large extent. If we can find that two base patterns are regularly found together in verbs, we can help the lexicographer by making suggestions while she/he is annotating the corpus. This study contributes to a more enriched and complex corpus annotation in which the system can help to find semantic regularities instead of a list of concordances with no inter-connection. This proposal could also be of help to providing clues for meaning differentiation and ordering in the dictionary entry (Jiang and Chen 2017).

2   Theoretical and Methodological Framework

The CPA patterns such as the ones shown in examples (1) and (2) are pieces of phraseology that are found frequently in a corpus. Each meaning of the verb is linked to one or more patterns of use. Observe the following example:

(4)   [[Human]] keeps [[Physical Object]] in [[Location]]

A pattern such as (4) is mapped to the meaning 'to store' of the verb *to keep*. This shows that the different meanings of the verb *to keep* can be disambiguated in context by analyzing the argument and syntactic structure of the verb and categorizing the arguments with semantic types. This theoretical and methodological line of research has its roots in a number of authors who observe that word meanings are disambiguated by context, e.g., Malinowski (1923), Firth (1935), Sinclair (1998) and Pustejovsky (1995), among others (see Hanks 2013 for a more detailed approach to the theoretical background of this technique). Firth (1935: 7) early states that "the complete meaning of a word is always contextual, and no study of meaning apart of a complete context can be taken seriously". According to the same author (Firth 1935: 7), this principle is what makes "systematic use of quotations or context" in dictionaries a crucial element for lexicographic representation of meaning. In the same way, CPA is a proposal for systematic corpus analysis of words, in which syntagmatic context of a verb in real discourse is analyzed and mapped into a meaning.
As already stated, verb patterns consist of basic valency structure, but an appropriate semantic categorization of each argument is also necessary. In (4), many words or phrases can be the subject of the verb (e.g., *student, you, Veronica, the new owners, we,* etc.), and all of them are unified under the same semantic type, [[Human]]. The same happens with [[Physical Object]] and [[Location]]. Semantic types are semantic categories that, in CPA, connect each other in an ontology of around 250 labels. Thus, while the verb *to keep* is ambiguous in isolation, patterns are unambiguous. In the present proposal, we pay special attention to semantic types and how they play a role in the configuration of verb meanings. For example, we can observe that pattern 1 and 2 in examples (1) and (2) have the same syntactic structure (transitive), but there is a variation in the semantic type of the direct object: in both examples (1) and (2), pattern 1 has [[Physical Object]] as direct object, while pattern 2 has [[Abstract Entity]]. This variation alone allows to differentiate the patterns, which are mapped onto different meanings.
Normal patterns such as (4) can be exploited for creative purposes or for fitting a specific communicative situation. For example, a sentence such as "There will be a large freezer for keeping food" can be considered a normal use of pattern (4), but a sentence such as "There will be a large freezer for keeping your pleasure" is not so common, but one can understand that there is a game of words in which the pleasure ([[Abstract Entity]]) a person gets by eating the food ([[Physical Object]]) stored in the freezer is materialized as something one can eat. Both norms and exploitations usually have their origin in a metaphor, and can be understood because, as stated in the introduction, metaphors are cognitive devices which are shared by the members of a community.
CPA is being used to build the *Pattern Dictionary of English Verbs* (Hanks, online) and the Verbario database for Spanish verbs (Renau, http://www.verbario.com). So far, Verbario contains 227 verbs, 1,233 patterns and 84,227 manually analysed concordances which are linked to the patterns. All CPA projects use the same method and ontology, which makes the data compatible (as was shown in Baisa et al. 2016) and allows to test the present proposal in other languages. For the purposes of the present study, the limitation of the technique lays on the fact that it is basically manual in spite of some attempts to

automatize certain parts of the task (Renau et al. 2019). Therefore, the analysis could be biased by the different annotators, and we do not possess data on a large-enough scale to generalize our results. Hence, the present study is explorative, and we expect to address these limitations in the future work.

3 Methodology

As already explained, we used Verbario's database as the source of our study units. For this preliminary study, we restricted our analysis to transitive patterns, that is, we included in our analysis only those verbs with 2 or more transitive patterns, with argument 1 as subject and argument 2 as direct object. We included in our study those patterns having a 3rd argument (e.g., adverbial or indirect object), but argument 3 was excluded from the study, because we were looking for identical syntactic structures in order to have semantic types as the only variable. We leave for future work to compare other possible structures and arguments, such as intransitive patterns *(Argument 1 + verb)* or trivalent structures with direct object or adverbial as argument 3. According to this, the first operation was to delete all the patterns with less than 2 transitive patterns, and for the remaining patterns, to keep the structure *Argument 1 + verb + Argument 2* and delete the rest of the pattern.

The second step was to transform the patterns into base patterns by deleting the verb and keeping the semantic types, for example:

(5)   *Pattern 1 of "ensanchar":* [[Humano]] ensanchar [[Objeto Físico]]   *Base pattern 1:* Humano ~ Objeto Físico
   Pattern 1 of "guardar": [[Humano]] guarda [[Objeto Físico]]

(6)   *Pattern 2 of "ensanchar":* [[Humano]] ensanchar [[Entidad Abstracta]]   *Base pattern 2:* Humano ~ Entidad Abstracta
   Pattern 2 of "guardar": [[Humano]] guarda [[Entidad Abstracta]]

(See (1), (2) and (3) for clarification.)

In case of semantic alternations–that is, when more than one semantic type alternate in the same argument (Hanks, 2013: 176-180), we split the verb pattern in each of the semantic types involved and created one base pattern for each semantic type. Observe the following example:

(7)   Verb *acarrear* 'to carry'

   Pattern 1: [[Humano | Vehículo]] acarrear [[Objeto Físico]] *([[Human | Vehicle]] carries [[Physical Object]])*

Vertical line between [[Human]] and [[Vehicle]] in example 7 means that both semantic types can be the subject of the verb in pattern 1 of *acarrear* 'to carry'. This alternation does not change the meaning of the pattern, which in this case is 'to take something somewhere'. In this respect, in example 8, we can observe a sentence in which we have [[Human]] as subject (*los niños* 'the children') and another one in which we have [[Physical Object]] as subject (*una caravana de camiones* 'a caravan of trucks'). Both sentences are linked to pattern 1 in example (7).

(8)   *Example (for [[Human]] as subject):* "*Los niños* enfrentan riesgos de seguridad y de salud al tirar y acarrear cargas pesadas" ('*Children* deal with security and health risks when pulling and carrying heavy loads').

   Example (for [[Vehicle]] as subject): "*Una caravana de camiones* se encarga de acarrear las provisiones" ('*A caravan of trucks* is responsible for carrying the provisions').

Hence, in our study, for the [[Human]] variant of the pattern, we created the base pattern Human ~ Physical Object, and for the [[Vehicle]] variant, we created the base pattern Vehicle ~ Physical Object.

Once the base patterns were extracted, we created a matrix where each column is a base pattern and each row a verb. We added the number of the pattern to the cell when there was a match (see table 1 for examples).

	Eventuality ~ State of Affairs	Eventuality ~ Cognitive State	Eventuality ~ Event	Eventuality ~ Human Group	Eventuality ~ Human	Physical Object ~ Physical Object
abrasar 'to burn'	0	0	0	0	3	2
abrigar 'to wrap up'	0	0	0	0	0	0
abrir 'to open'	0	16	0	0	0	0
acarrear 'to carry'	3	0	0	0	0	0
aconsejar 'to advise'	0	0	0	0	0	0
acortar 'to shorten'	0	0	0	0	0	0
acosar 'to harass'	0	0	0	0	3	0

acostar 'to lay down'	0	0	0	0	0	0
activar ' to activate'	0	0	0	0	0	0
agrietar 'to crack'					2	1

Table 1: Fragment of the matrix for illustration.

In table 1 we observe how, in this fragment of the matrix, there are 7 matches between verbs and base patterns. For example, pattern 16 of *abrir* 'to open' has EVENTUALITY ~ COGNITIVE STATE as verb pattern.

After this operation, we calculated the frequency in which a base pattern X appeared together with base pattern Y in the same verb, and in how many verbs we have the same coincidence. For example, in table 1, *abrasar* and *acosar* share the base pattern EVENTUALITY ~ HUMAN, but these verbs do not share any other pattern –nor in table 1, which is a fragment for illustration, nor in any other part of the whole matrix. This means that this combination of two verbs is not a candidate to find possible pairs such as the ones shown in (7) and (8). Conversely, *abrasar* and *agrietar* 'to crack' do share two combinations: PHYSICAL OBJECT ~ PHYSICAL OBJECT and EVENTUALITY ~ HUMAN. These combinations are the target of our study. In this case, for example, these pairs of base patterns exhibit a metaphorical relation in which [[Eventuality]] behaves with [[Human]] the same way as [[Physical Object]] with another [[Physical Object]]. Table 2 shows this example in more detail: we observe how the event of 'burning something physical' (in the case of *abrasar* 'to burn') is transferred to 'emotionally burning a person'. In parallel, the event of 'cracking something physical' (in *agrietar* 'to crack') is transferred to 'morally cracking a person'. The same metaphor underlies both verbs in the same way.

Base pattern	Verb	Pattern	Implicature	Example
PHYSICAL OBJECT ~ PHYSICAL OBJECT	*abrasar* 'to burn'	*Pattern* 2 [[Objeto Físico]] abrasar [[Objeto Físico]] *([[Physical Object]] burns [[Physical Object]])*	*[[Physical Object]] makes that [[Physical Object]] is very hot.*	…los hierros abrasando la carne. *(...the irons burning the flesh.)*
	agrietar 'to crack'	*Pattern 1* [[Objeto Físico]] agrietar [[Objeto Físico]] *([[Physical Object]] cracks [[Physical Object]])*	*[[Physical Object]] makes that cracks appear in [[Physical Object]].*	Una bola de granizo agrieta tu cristal. *(A hail ball cracks your glass.)*
EVENTUALITY ~ HUMAN	*abrasar* 'to burn'	*Pattern 3* [[Eventualidad]] abrasar [[Humano]] *([[Eventuality]] burns [[Human]])*	*[[Eventuality]] has an strong and negative effect on [[Human]].*	Seguir pensando [en que] no sé qué hacer con mi vida me abrasaba. *(Continuing thinking that I do not know what to do with my life burned me.)*
	agrietar 'to crack'	*Pattern 2* [[Eventualidad]] agrietar [[Humano]] *([[Eventuality]] cracks [[Human]])*	*[[Eventuality]] weakens [[Human]], makes her/him lose her/his power or strength.*	Esa tarde, otra vez lo agrietó el descreimiento. *(That afternoon, disbelief cracked him again.)*

Table 2: An example of a match of base patterns in two verbs: *abrasar* 'tu burn' and *agrietar* 'to crack' (see table 1, green cells). The implicatures are paraphrases of the patterns which explain their meanings (Hanks, 2013).

Finally, we also calculated an association coefficient which indicated the grade of reciprocity in which a pair of base patterns appears in the same verb. To do this, we applied the following formula:

$$\frac{f(i\,j)}{\sqrt{f(i)} \cdot \sqrt{f(j)}}$$

i and *j* are 2 base patterns appearing in the same verb. Frequency of *i* and *j* appearing together is divided by the total frequency of *i* per the total frequency of *j*. Square root is used to mitigate the difference between the highest and lowest numbers. Numbers were multiplied by 100 in order to avoid decimals.

Results of frequency and association were displayed in a table.

4   Results and Discussion

Of the total 227 verbs in the database, we found 177 with 2 or more transitive patterns (78% of the verbs), that is to say, 78% of the verbs were included in the study (in the matrix, they were displayed in the rows, as shown in table 1). Of these verbs, we obtained 510 base patterns (in the matrix, they were displayed in the columns, as shown in table 1). 32 of the base patterns (6,27%) associate with another base pattern creating 77 pairs. These pairs appear in 111 of the 177 verbs (62,7%). This means that only a few base patterns combine with another base pattern two or more times, but at least one of these combinations can be traced in most verbs. For example, the combination HUMAN ~ EMOTION → HUMAN ~ HUMAN is very frequent (n = 5, 71%): thus, in 71% of the verbs in which we find the base pattern HUMAN ~ EMOTION, we also find the base pattern HUMAN ~ HUMAN. Combinations are not commutative, for example, the combination HUMAN ~ HUMAN → HUMAN ~ EMOTION is very rare: only in 5% (n = 5) of the verbs in which we find the base pattern HUMAN ~ HUMAN the base pattern HUMAN ~ EMOTION is present too. This is normal, as HUMAN ~ HUMAN is much more frequent per se than HUMAN ~ EMOTION.

Figure 1 shows the semantic types which were found more frequently (> 10) in the base patterns of the data sample:
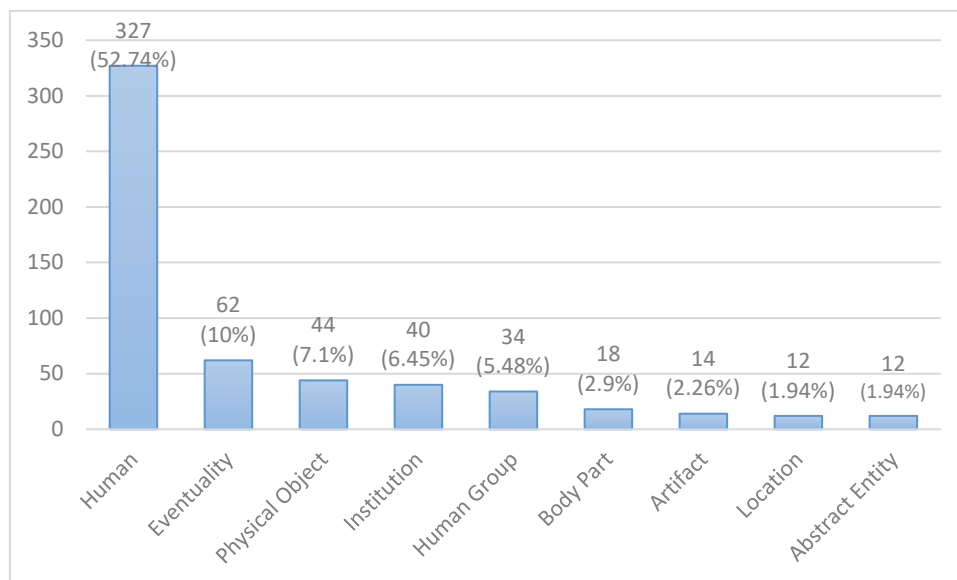

Figure 1: The most frequent semantic types taking part in recurrent combinations of base patterns.

It is not surprising to find these semantic types in figure 1 because they are common categories found in general in CPA patterns. In addition, we can observe that they take part in the formation of many metaphors, e.g. "Humans are Artifacts", "Abstract Entity is Physical Object", "Physical Object is Human", etc.

Figure 2 shows the most frequent (> 5) base patterns found in a recurrent association with another base pattern:
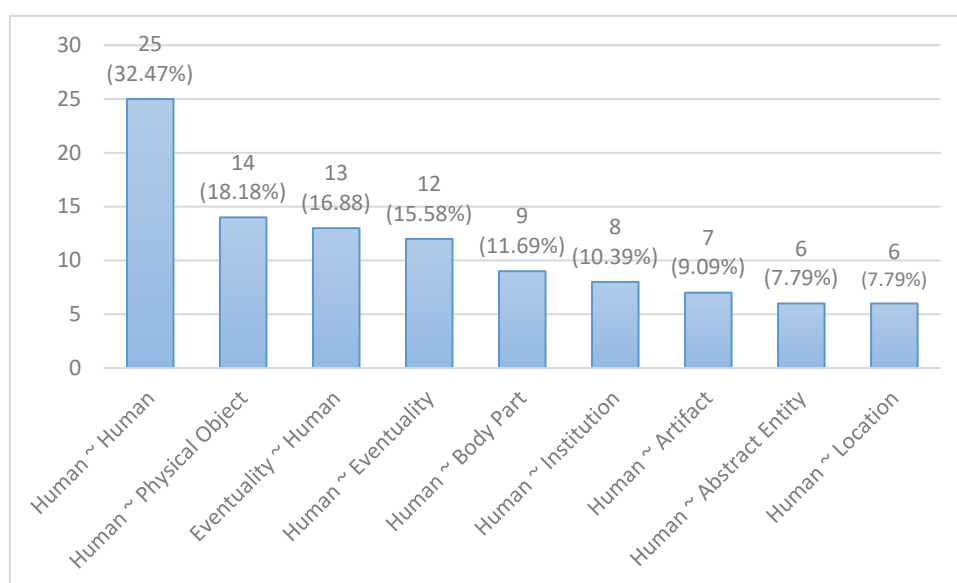

Figure 2: The most frequent base patterns found in association with other base patterns in the data sample.

Considering the data shown in figure 2, it is normal to find HUMAN as a frequent semantic type in table 2 as well. We find

it usually as the subject of the sentence, probably as agent of the event which is directed to abstract or concrete entities, or eventualities.

Table 3 shows the most frequent (> 5) associations of base patterns in the sample:

Base pattern 1 →	Base pattern 2	Verbs (n)	Verbs (%)
HUMAN ~ HUMAN	HUMAN ~ PHYSICAL OBJECT	19	10,73
HUMAN ~ EVENTUALITY	HUMAN ~ HUMAN	14	7,91
EVENTUALITY ~ HUMAN	HUMAN ~ HUMAN	11	6,21
HUMAN ~ BODY PART	HUMAN ~ HUMAN	11	6,21
HUMAN ~ ARTIFACT	HUMAN ~ HUMAN	10	5,65
HUMAN ~ ABSTRACT ENTITY	HUMAN ~ PHYSICAL OBJECT	9	5,08
HUMAN ~ STATE OF AFFAIRS	HUMAN ~ HUMAN	9	5,08
HUMAN ~ EVENT	HUMAN ~ HUMAN	9	5,08
HUMAN ~ HUMAN	HUMAN ~ LOCATION	9	5,08
HUMAN ~ HUMAN	PHYSICAL OBJECT ~ PHYSICAL OBJECT	9	5,08

Table 3: The most frequent associations of base patterns in the sample. Percentages are in relation with the total 177 verbs of the sample.

These results do not necessarily show that the association has a metaphorical nature, but coincide with types of metaphors which have been reported in the literature, and we could hypothesize that they are clues to identify metaphors. For example, in the pair HUMAN ~ HUMAN → HUMAN PHYSICAL OBJECT, the underlying metaphor could be "Humans are Physical Objects". The logic under these associations would be that certain events (represented by the verb) could be the source domain for other types of events which, as target domain, are characterized as having certain similarities with the source domain. In section 5 we will present a case study to try to show this rationale.

Table 4 shows the frequency in which a base pattern appears in a verb together with another base pattern:

Base pattern 1 →	Base pattern 2	Verbs (n)	Verbs (%)
EVENTUALITY ~ EVENTUALITY	HUMAN ~ HUMAN	6	100
HUMAN ~ ANYTHING	HUMAN ~EVENTUALITY	6	100
PHYSICAL OBJECT ~ HUMAN	HUMAN ~ HUMAN	5	100
EVENTUALITY ~ PSYCHOLOGICAL TRAIT	HUMAN ~ HUMAN	4	100
INSTITUTION ~ EVENT	HUMAN ~ HUMAN	4	100
HUMAN ~ STUFF	HUMAN ~ HUMAN	7	85
EVENTUALITY ~ EVENTUALITY	EVENTUALITY ~ HUMAN	5	80
HUMAN ~ EMOTION	HUMAN ~ HUMAN	7	71
HUMAN ~ ANYTHING	HUMAN ~ HUMAN	6	66
HUMAN ~ INFORMATION	HUMAN ~ HUMAN	11	63
HUMAN GROUP ~ HUMAN	INSTITUTION ~ INSTITUTION	7	57
HUMAN GROUP ~ HUMAN	INSTITUTION ~ HUMAN	7	57
HUMAN ~ STUFF	HUMAN ~ ABSTRACT ENTITY	7	57
HUMAN GROUP ~ HUMAN	HUMAN ~ INSTITUTION	7	57
HUMAN ~ STUFF	EVENTUALITY ~ HUMAN	7	57
HUMAN ~ ABSTRACT ENTITY	HUMAN ~ PHYSICAL OBJECT	17	52
HUMAN ~ STATE OF AFFAIRS	HUMAN ~ HUMAN	17	52

Table 4: Frequency of combinations of base patterns (> 50% verbs).

Table 4 shows how in 17 of the 77 combinations of patterns (22,07%) we can find a frequent combination. In particular, in 5 of the verbs (6,5%) the association between base patterns 1 and 2 covers 100% of the cases, that is, each time that we find base pattern 1, we also find base pattern 2. These results show that base patterns in column 1 are good predictors for the existence of base patterns in column 2 in the data sample. If these results were corroborated by studies with larger data, the method would be appropriate as an assistance to detect new meanings in corpus.

Finally, table 5 shows the highest (> 30) association coefficients of the pairs.

Base pattern 1 →	Base pattern 2	Verbs (n)	Coefficient association	Verbs (%)
HUMAN GROUP ~ HUMAN	INSTITUTION ~ INSTITUTION	4	47	57%
INSTITUTION ~ INSTITUTION	HUMAN ~ HUMAN	4	47	40%
HUMAN GROUP ~ HUMAN GROUP	INSTITUTION ~ INSTITUTION	4	44	50%
INSTITUTION ~ INSTITUTION	HUMAN GROUP ~ HUMAN GROUP	4	44	40%
HUMAN ~ ANYTHING	HUMAN ~ EVENTUALITY	6	41	100%
HUMAN ~ EVENTUALITY	HUMAN ~ ANYTHING	6	41	17%
HUMAN GROUP ~ HUMAN	INSTITUTION ~ HUMAN	4	39	57%
INSTITUTION ~ HUMAN	HUMAN GROUP ~ HUMAN	4	39	26%
HUMAN ~ STUFF	HUMAN ~ ABSTRACT ENTITY	4	36	57%
HUMAN GROUP ~ HUMAN GROUP	INSTITUTION ~ HUMAN	4	36	50%
INSTITUTION ~ HUMAN	HUMAN GROUP ~ HUMAN GROUP	4	36	26%
HUMAN ~ ABSTRACT ENTITY	HUMAN ~ STUFF	4	36	23%
EVENTUALITY ~ EVENTUALITY	EVENTUALITY ~ HUMAN	4	32	80%
HUMAN ~ ABSTRACT ENTITY	HUMAN ~ PHYSICAL OBJECT	9	32	52%
HUMAN ~ PHYSICAL OBJECT	HUMAN ~ ABSTRACT ENTITY	9	32	20%
EVENTUALITY ~ HUMAN	EVENTUALITY ~ EVENTUALITY	4	32	12%
INSTITUTION ~ INSTITUTION	HUMAN ~ ARTIFACT	5	31	50%
HUMAN ~ PHYSICAL OBJECT	HUMAN ~ HUMAN	19	31	43%
HUMAN ~ HUMAN	HUMAN ~ PHYSICAL OBJECT	19	31	22%
HUMAN ~ ARTIFACT	INSTITUTION ~ INSTITUTION	5	31	19%

Table 5: Coefficient association > 30 for the base pattern pairs. Frequency is shown for reference.

Table 5 shows that these pairs exhibit a strong association. The association is not reciprocal, though: for example, while the base pattern HUMAN ~ ANYTHING appears together with HUMAN ~ EVENTUALITY in the same verb 100% of the times, it is not as frequent that HUMAN ~ EVENTUALITY appears together with HUMAN ~ ANYTHING.

5   Case Studies

In this section, we give a more detailed description of results regarding some of the pairs of base patterns in the data sample. We want to observe if, as we stated in section 1, we could find potential sources for metaphors. Table 6 shows results for the pair HUMAN ~ HUMAN / PHYSICAL OBJECT ~ HUMAN.

Base pattern 1 →	Base pattern 2	Verbs (n)	Association coefficient	Frequency	Verbs
HUMAN ~ HUMAN	PHYSICAL OBJECT ~ HUMAN	5	24	5%	*aplastar, cubrir, dañar, estorbar, estremecer* ('to crush, to cover, to harm, to hinder, to shake')
PHYSICAL OBJECT ~ HUMAN	HUMAN ~ HUMAN			100%	

Table 6: Frequency and coefficient association for the pairs HUMAN ~ HUMAN → PHYSICAL OBJECT ~ HUMAN and for PHYSICAL OBJECT ~ HUMAN → HUMAN ~ HUMAN.

Table 6 shows that the association HUMAN ~ HUMAN → PHYSICAL OBJECT ~ HUMAN is very infrequent (5%), while the association PHYSICAL OBJECT ~ HUMAN → HUMAN ~ HUMAN takes place 100% of the times. The verb patterns are the following ones:

(9)   Verb *aplastar* 'to crush'
    Pattern 1: [[Objeto Físico]] aplasta a [[Humano]] *([[Physical Object]] crushes [[Human]])*
    Example: "Una mujer de 54 años falleció ayer aplastada por un vehículo de limpieza" *('A 54-year-old woman died yesterday crushed by a cleaning vehicle').*

    Pattern 2: [[Humano]] aplasta a [[Humano]] *([[Human]] crushes [[Human]])*
    Example: "Esta joven de progresión imparable aplasta a sus rivales sin compasión" *('This young woman of*

unstoppable progression crushes her rivals without compassion').

(10) Verb *cubrir* 'to cover'
Pattern 1: [[Objeto Físico]] cubre a [[Humano]] *([[Physical Object]] covers [[Human]])*

Example: "Detestan las pieles que las cubren [a las mujeres ricas]" *('They hate the furs that cover them [the rich women]').*

Pattern 2: [[Humano]] cubre a [[Humano]] *([[Human]] covers [[Human]])*
Example: "Tengo miedo de que todavía se estén cubriendo unos a otros" *('I am afraid they are still covering for each other').*

(11) Verb *dañar* 'to harm'
Pattern 1: [[Objeto Físico]] daña a [[Humano]] *([[Physical Object]] harms [[Human]])*
Example: "Esta potente escopeta puede dañar a varios enemigos" *('This powerful shotgun can harm several enemies').*

Pattern 2: [[Humano]] daña a [[Humano]] *([[Human]] harms [[Human]])*
Example: "Sabe expresar sus emociones como las siente sin dañar a los demás" *('She/he knows how to express her/his emotions without harming others').*

(12) Verb *estorbar* 'to hinder, to disturb'
Pattern 1: [[Objeto Físico]] estorba a [[Humano]] *([[Physical Object]] hinders [[Human]])*
Example: "Esquivaba los pocos autos que le estorbaban en el camino" ('She/he dodged the few cars that stood in her/his way').

Pattern 2: [[Humano]] estorba a [[Humano]] *([[Human]] bothers [[Human]])*
Example: "Me dirigí al jardín para no estorbar a los adultos" *('I went to the garden to not disturb the adults').*

(13) Verb *estremecer* 'to shake'
Pattern 1: [[Objeto Físico]] estremece a [[Humano]] *([[Physical Object]] makes [[Human]] shake)*
Example: "Ese cuadro la estremece como ninguna otra cosa" *('That painting makes her shake like nothing else').*

Pattern 2: [[Humano]] estremece a [[Humano]] *([[Human]] harms [[Human]])*
Example: "El predicador la estremecía con sus emociones personales" *('The preacher shook her with his personal emotions').*

All examples (9) to (13) exhibit a metaphorical relation between base pattern 1 and 2. Metaphors are based on the categorization of [[Humans]] as [[Physical Objects]] that can make actions to other [[Humans]] which are similar to the ones that [[Physical Objects]] can do to [[Humans]]. We could formalize the metaphors underlying patterns (9) to (13) as follows:

- (9) "Somebody morally crushing somebody is an object physically crushing her/him".
- (10) "Somebody covering for somebody is an object covering her/him".
- (11) "Somebody morally harming somebody is an object physically harming somebody".
- (12) "Somebody disturbing somebody is an object physically disturbing her/him".
- (13) "Somebody emotionally shaking somebody is an object physically shaking her/him".

As we can observe, this formulation is rich in information and explain the semantic and cognitive relation between two meanings of a verb, and between different verbs. All of them share the basic idea that events caused by a [[Physical Object]] and experimented by a [[Human]] can be used to understand moral or emotional events. These preliminary findings are promising in the sense that they could be empirical linguistic data to corroborate the theory of conceptual metaphor.
It is interesting, though, that there are cases in which the two base patterns do not have a metaphorical relation, because both are metaphors of another base pattern. Observe the following example taken from the pair HUMAN ~ HUMAN / HUMAN ~ PHYSICAL OBJECT:

(14) Verb *comer* 'to eat, win, crash'

Pattern 1: [[Humano]] (se) come a [[Humano]] *([[Human]] beats [[Human]])*
Example: "No podemos dejar que nos coman, hemos de imponernos" *('We cannot let them beat us, we have to impose ourselves').*

Base pattern 1: HUMAN ~ HUMAN

Pattern 2: [[Humano]] (se) come [[Objeto Físico]] *([[Human]] crashes against [[Physical Object]])*
Example: "Fidel se comió un escenario a raíz de un traspié" *('Fidel crashed against a stage as a result of a stumble')*

Base pattern 2: HUMAN ~ PHYSICAL OBJECT

In (14), both patterns, 1 and 2, have their origin in the pattern *[[Humano | Animal]] (se) come [[Comida]] ([[Human | Animal]] eats [[Food]]),* which is the most common, literal meaning. This pattern originates different figurative meanings of the verb, such as the ones shown in (14), which take the act of 'eating food' to refer to the act of 'clearly winning somebody in a competition' (pattern 1) (as 'devouring somebody'), or 'violently contacting something' (a humorous way of describing a crash when a person collides head-on with an object) (pattern 2). Hence, there is no metaphorical relation between base patterns 1 and 2, but between these two patterns and the one mapped to the literal meaning 'to eat'.
Finally, another aspect of the qualitative analysis of the data is the fact that the method cannot predict the direction of the metaphor. In all cases from (9) to (13) the source domain is the base pattern PHYSICAL OBJECT ~ HUMAN, and HUMAN ~ HUMAN is the target domain. However, observe the following examples of the pair HUMAN ~ HUMAN / HUMAN ~ PHYSICAL OBJECT:

(15) Verb *albergar* 'to give accommodation, to store'

Pattern 1: [[Humano]] alberga a [[Humano]] *([[Human]] gives accommodation to [[Human]])*
Example: "...venteros socarrones como el que alberga a Don Quijote y Sancho" *('...sardonic inn owners such as the one who gives accommodation to Don Quijote and Sancho').*

Base pattern 1: HUMAN ~ HUMAN

Pattern 2: [[Humano]] alberga [[Objeto Físico]] *([[Human]] stores [[Physical Object]])*
Example: "En su casa albergaba una colección de libros" *('At her/his home she/he stored a book collection').*

Base pattern 2: HUMAN ~ PHYSICAL OBJECT

(16) Verb *cortar* 'to cut, to interrupt'

Pattern 1: [[Humano]] cortar [[Objeto Físico]] *([[Human]] cuts [[Physical Object]])*
Example: "Seguí cortando la leña" *('I kept cutting the firewood').*

Base pattern 2: HUMAN ~ PHYSICAL OBJECT

Pattern 2: [[Humano]] corta a [[Humano]] *([[Human]] interrupts [[Human]])*
Example: "Me cortó con una sequedad que me dejó desorientado" *('She/he cut me with such a brusqueness that left me disoriented').*

Base pattern 2: HUMAN ~ HUMAN

We can observe that, in (15), HUMAN ~ HUMAN is the source domain for HUMAN ~ PHYSICAL OBJECT, as [[Physical Objects]] are categorized as such valuable things that are like persons that one hosts in a place. In contrast, in (16), the source domain is HUMAN ~ PHYSICAL OBJECT and the target domain is HUMAN ~ HUMAN, as conversation is categorized as something that can be "cut" or interrupted as a [[Physical Object]]. These findings, together with the ones exemplified in (14), lead us to believe that the method could be used to automatically detect these association of base patterns in corpus, but it would be necessary for human analysis to corroborate that the associations are really metaphors.

6 Conclusions and Future Work

This proposal is a preliminary attempt to find regularities in combinations of verb patterns with the same syntactic structure, using semantic types only. The purpose was to find out if these regularities have a metaphorical origin. The method is very simple and it can be used in all CPA projects or similar types of corpus annotation. Results seem promising, but they have to be tested with more data. It is also necessary to have a better theoretical articulation between types of metaphors and our corpus-driven findings, which could allow us to refine our results.
This study leads us to the following preliminary conclusions:

- Patterns of usage of different verbs share common features, particularly when focusing on semantic types. This common, general semantic information, that we called *base patterns,* can be easily extracted from verb patterns and traced in all verbs.
- A small number of base patterns form combinations which cover most of the verbs in the sample, which leads to consider that the proposed method could be appropriate to find new patterns of usage in corpus, linked to new meanings, and to organize the lexicographic information in the entry.

- Results also show that a small number of semantic types cover most of the base patterns, and a small number of base patterns connect to create pairs. This Zipfian tendency allows us to consider that further studies with this small group of semantic types and pairs could be cost-effective and cover many other cases.

- We do not have enough data to confirm that this method is appropriate to find metaphors in corpus, but the preliminary results show that in many cases we find a metaphorical origin in the connections between base patterns. Thus, at the moment, results show that base patterns linked via metaphors are common.

- While, certainly, most of the analysed pairs exhibit a metaphorical relation, the source domain of the metaphor is not predictable, because many semantic types can work as source or target domain. This is not new in the theory of metaphor, but the present analysis provides us with types of source domain / target domain combinations that may not be very frequent or prototypical. These data could bring new insights regarding the typology of metaphors.

- Similarly, combinations of base patterns provide us with information which could potentially enrich the usual formulation of conceptual metaphors. For example, a usual metaphor is "Events are Physical Objects", such as in "Economic crisis hit the industry" (like a hammer), "Their relationship cracked" (like a wall), etc. Our results could add information to this formulation, such as we proposed in section 5, e.g. "An event cracking is an object physically cracking".

There are different lines for future research. The most important next task would be to apply the same method to a more extensive group of verbs and see if results were consistent with the present ones. Another possible line of work could be to replicate the same procedure by converting semantically specific semantic types into more general labels, e.g. converting [[Illness]] into [[Eventuality]], [[Emotion]] into [[Abstract Entity]], etc. This can be done because semantic types are organized in an ontology and linked to each other via IS-A relations. This way it is possible to find more recurrent, general metaphors and this would allow us to establish a corpus-driven taxonomy of metaphors. This operation can be easily automatized since we already have a machine-readable version of the CPA Ontology. Finally, as already mentioned, the procedure can be replicated in other CPA projects such as the *Pattern Dictionary of English Verbs* (Hanks, online).

7   References

Apresjan, J. D. (1974). Regular Polysemy. In *Linguistics,* 12(142), pp. 5-32.

Atkins, B.T.S. & Rundell, M. (2008). *The Oxford Guide to Practical Lexicography.* Oxford: Oxford University Press.

Baisa, V., Može, S. & Renau, I. (2016). Multilingual CPA: Linking Verb Patterns across Languages. In T. Margalitadze, G. Meladze (eds.), *Proceedings of the XVII Euralex International Congress.* Tbilisi: Tbilisi University Press, pp. 410-417.

Firth, J. R. (1935[1957]). The Technique of Semantics. In J. R. Firth, *Papers in Linguistics.* Oxford: Oxford University Press, pp. 7-33.

Hanks, P. (2004). Corpus Pattern Analysis. In G. Williams, S. Vessier (eds.), *11th Euralex International Congress. Proceedings.* Lorient: Université de Bretagne-Sud, pp. 87-97.

Hanks, P. (2013). *Lexical Analysis: Norms and Exploitations.* Cambridge, Ma: MIT Press.

Hanks, P. (dir.). (Online). *Pattern Dictionary of English Verbs.* Accessed at: http://www.pdev.org.uk

Hartmann, R.R.K. (2001). *Teaching and Researching Lexicography.* Harlow, UK: Pearson Education.

Jiang, G. & Chen, Q. (2017). A Micro Exploration into Learner's Dictionaries: A Prototype Theoretical Perspective. In *International Journal of Lexicography,* 30(1), pp.108-139.

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P. & Suchomel, V. (2014). The sketch engine: ten years on. In *Lexicography*, 1(1), pp. 7-36.

Kosem, I. (2016). Interrogating a Corpus. *The Oxford Handbook of Lexicography.* Oxford: Oxford University Press, pp. 76-93.

Lakoff, G. & Johnson, M. (1980). *Metaphors We Live By.* Chicago: The University of Chicago Press.

Malinowski, B. (1923). The Problem of Meaning in Primitive Languages. In C. K. Odgen I. A. Richards (eds.). *The Meaning of Meaning: A Study of the Influence of Language upon Thought and of the Science of Symbolism.* Cambridge: Cambridge University Press, pp. 296-336.

Moldovan, A. (2019). Descriptions and Tests for Polysemy. In *Axiomathes*, *31(3),* pp. 1-21.

Renau, I. (2021). Algunos datos lexicográficos y de corpus para la representación de la polisemia regular en los diccionarios. *Boletín de Filología,* Anexo, pp. 905-925.

Renau, I., Nazar, R., Castro, A., López, B. & Obreque, J. (2019). Verbo y contexto de uso: un análisis basado en corpus con métodos cualitativos y cuantitativos. In *Revista Signos,* 52(101), pp. 878-901.

Renau, I. (dir.). (Online). *Verbario.* Accessed at: http://www.verbario.com

Pustejovsky, J. (1995). *The Generative Lexicon.* Cambridge, Ma: MIT Press.

Sinclair, J. (1998[2004]). The Lexical Item. In J. Sinclair, *Trust the Text. Language, Corpus and Discourse.* Amsterdam: Routledge, pp. 131-148.